

## APPLICATIONS OF ARTIFICIAL INTELLIGENCE FOR CHEMICAL INFERENCE—V\*

### AN APPROACH TO THE COMPUTER GENERATION OF CYCLIC STRUCTURES. DIFFERENTIATION BETWEEN ALL THE POSSIBLE ISOMERIC KETONES OF COMPOSITION $C_6H_{10}O$ .

YOUNUS M. SHEIKH, ARMAND BUCHS,† ALLAN B. DELFINO,‡ GUSTAV SCHROLL,§  
A. M. DUFFIELD, CARL DJERASSI, B. G. BUCHANAN, G. L. SUTHERLAND,  
E. A. FEIGENBAUM and J. LEDERBERG

Departments of Chemistry, Computer Science and Genetics, Stanford University,  
Stanford, California 94305, USA

(Received 1 July 1970; accepted (revised) 7 August 1970)

**Abstract**—The computer program DENDRAL has been modified so as to include cyclic structures for the first time. As a result a list of all the possible isomers (linear and cyclic) of selected compositions can now be generated. The number of cyclic structures exceeds that of the linear molecules for a given composition. A method, based on their physical properties (i.r., n.m.r. and mass spectra), for the identification of each of the 27 possible ketones (exclusive of 5 cyclopropanones) of composition  $C_6H_{10}O$  is described.

RECENT publications<sup>1,2,3</sup> from our interdisciplinary research group describe the development and refinement of a computer program (Heuristic DENDRAL) capable of the structural elucidation of unknown saturated aliphatic ketones,<sup>2</sup> ethers<sup>3</sup> and amines<sup>1</sup> from their low resolution mass spectra. While we readily admit that the chosen functional groups and their molecular environments represent only a small area of interest to an organic chemist, nonetheless these papers represent a start at direct computer interpretation of physical data without human intervention. The physical data used to date have primarily been low resolution mass spectra, and when available, n.m.r. spectra. Our experience is that the correct structure always appeared in the final output. This output list may consist of only one or a few selected out of many thousands, often millions, of possible structures.

The present report is a significant extension of the domain of candidate structures, namely cyclic molecules. The variety and complexity of ring-containing structures of course gives them a major place in the work of the organic chemist. They also introduce many problems of symmetry beyond those seen in acyclics, and indeed the efficient generation, without overwhelming redundancy, of exhaustive lists of isomeric structures is a major challenge to the Heuristic DENDRAL STRUCTURE GENERATOR,|| the program module responsible for fetching (in principle) all possible candidate molecules as potential solutions to an analytical problem.

A completely general approach to the design of a 'ring-generator' has been formulated in some detail.<sup>4</sup> It is based on a major classification of rings according to

\* For Part IV, see Ref. 1.

† On leave of absence from the University of Geneva, Geneva, Switzerland.

‡ Present address: Allen-Babcock Computing, Palo Alto, California, USA.

§ Recipient of a Fulbright Travel Award. Present Address: Chemical Laboratory II, University of Copenhagen, Denmark.

|| Program modules are written in upper case.

the paths between the vertex atoms, namely those atoms at which rings are 'fused' in conventional descriptions. This corresponds to the set of trivalent, cyclic graphs, the enumeration of which has been programmed, computed and described.<sup>5</sup> Writing the complete program for all possible ring molecules should offer no special difficulties except tedium and running time and we have deferred doing this in favor of more general problems that might encompass this almost incidentally. (For example, we should write programs that 'understand' group theory, and can save us the tedium of transmitting an infinity of small tricks and insights for weeding out symmetries *prospectively*, as indeed was done in producing DENDRAL up to this point.)

The perfectly general case may also have limited utility compared to applications where some constraints on the character of possible rings are implicit in a problem statement.

At this time, we deal with a specific problem (ketonic isomers of  $C_6H_{10}O$ ) which encompasses, at most, a single, simple ring. We have then described all the relevant rings through a special purpose program. In effect, this takes all possible pairwise selections of 2 hydrogen atoms from acyclic isomers of  $C_6H_{12}O$  ( $2H + C_6H_{10}O$ ), and replaces the pair with a cross-link, in effect closing a ring. Some graph theoretical tricks are introduced to minimize redundancies, and retrospective checks for isomorphism were also inserted to eliminate them. This approach generates the entire molecule, including side chains. The potential redundancy approaches the number of ways that the ring-molecule can be cut to produce different trees.

A second approach, closer to the spirit of generalized DENDRAL, regards every pure ring (i.e. no side chains) as a SUPERATOM. A full list of these superatoms is generated by the function MONOCYCLIC. Each SUPERATOM (RING) is then embellished with the appropriate permutations of side chains, account being taken of the symmetry of the RING.

Although we have described<sup>6</sup> the computer program, DENDRAL, which generated complete and irredundant lists of the number of acyclic isomers which can be formulated from a given composition, no computer routine then existed for the enumeration of cyclic molecules. In view of the substantial increase in the number of isomers (and thus the cost of computer time) of a selected composition when cyclic representations are generated, it has only been feasible to produce the numbers of possible isomers for comparatively small empirical formulac (Tables 1 and 2). For example it takes approximately 6.2 seconds to generate all the non cyclic isomers of  $C_4H_6O$  and 20.3 seconds for the non-cyclic isomers of  $C_5H_8O$ ; but it takes approximately 14.7 seconds for all the cyclic isomers of  $C_4H_6O$  and 66.0 seconds for all the cyclic isomers of  $C_5H_8O$ . Work with larger compositions will have to wait until an

TABLE 1. NON-CYCLIC STRUCTURES OF COMPOSITION  $C_nH_{2n-2}O$

<i>n</i>	Ketones	Aldehydes	Ketenes	Alcohols	Ethers	Total
2	0	0	1	0	0	1
3	0	1	1	1	1	4
4	1	3	2	4	5	15
5	4	8	3	14	17	46
6	13	21	7	47	62	150
7	40	56	13	182	207	498

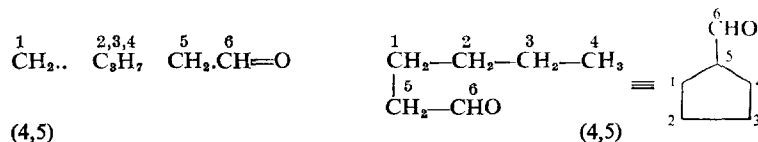
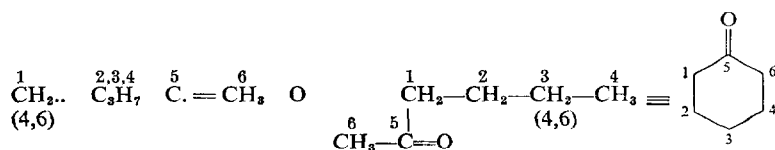
TABLE 2. CYCLIC STRUCTURES OF COMPOSITION  $C_nH_{2n-2}O$ 

$n$	Ketones	Aldehydes	Alcohols	Ethers	Total	Acyclic + Cyclic
2	0	0	0	0	0	1
3	1	0	2	4	7	11
4	2	1	10	20	33	48
5	7	4	55	79	145	191
6	19	13	215	310	558	708
7	57	47	???	???	???	???

alternative algorithm is developed. The following examples with their familiar chemical representations are presented as illustrations of the representation of cyclic structures in DENDRAL dot notation.<sup>6</sup>

The program derives its *canonical* notation from the following three points:

1. A structure is built from the lowest valued DENDRAL acyclic form.
2. The skeletal atoms are numbered in the order of their appearance in the dot notation.
3. Ring construction is achieved by using the smallest valued pair of numbers to describe a given cross-link.



Three modes exist for the routine operation of the DENDRAL computer program. First it can construct all the topologically possible acyclic isomers; second, using the option of a BADLIST\* filter, the chemically stable isomers can be selected; or third, using the GOODLIST constraints, only those isomers containing a specific structural unit (e.g. ketones, alcohols, etc.) will appear. The concept of a BADLIST for ring DENDRAL is more complex than for linear structures. In the latter instances organic chemists will readily agree on those groups of atoms considered to be chemically unstable; however, when these groups of atoms are embedded in ring structures they often will be stable. Furthermore, steric considerations must be considered in the implementation of ring DENDRAL. For instance the following restrictions have been placed on BADLIST: no isomers are accepted with a double bond in a three-membered ring (since these compounds would only be of limited stability); no triple

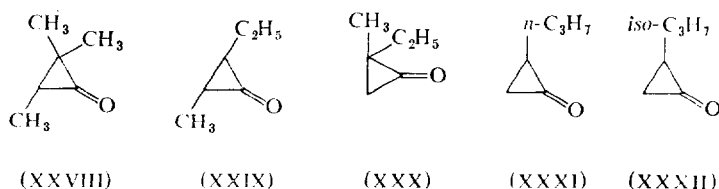
\* BADLIST, described in previous publications,<sup>2,3,6</sup> is the list of substructures which must not become embedded in molecular structures because, for example, they result in unstable molecules. Similarly, GOODLIST is the list of substructures which the program must build into every molecule it generates.

bond can be inserted in less than an eight-membered ring; and allenes are prohibited unless the ring size exceeds 8 atoms.

Table 1 is a new listing of the number of acyclic isomers of the formula  $C_nH_{2n-2}O$  as generated by the computer program DENDRAL. The number of ketones, aldehydes, ketenes, alcohols and ethers of a given composition were generated by placing the appropriate subgraph on GOODLIST.

Table 2 contains the number of possible cyclic structures of the same composition  $C_nH_{2n-2}O$  classified according to the number of ketones, aldehydes, alcohols and ethers. It is apparent that once ring systems are considered, the number of topologically possible isomers of these groups for any composition exceeds the number of their linear counterparts. Furthermore, while for  $C_6H_{10}O$  the number of cyclic ketones and cyclic aldehydes remains relatively low (19 and 13 respectively) the number of cyclic alcohols (215) and ethers (310) constitutes a large majority of the total structures.

It seemed desirable at this stage to confront Heuristic DENDRAL (the computer program) with the task of differentiating between each isomer of a given composition by the physical data (mass, n.m.r., i.r. and u.v. spectrometry) available for each isomer. In view of the scope of the synthetic problem (it being realized that many, probably the majority, of the isomers listed are as yet unknown and would have to be synthesized) we concentrated our attention on the question of differentiating between each of 27 possible ketones of composition  $C_6H_{10}O$ . It was decided to omit the five cyclopropanones (XXVIII to XXXII) as these compounds would be of limited stability.



The location of the carbonyl absorption in the infrared spectra (Table 3) of the 27 isomeric ketones (I to XXVII) serves as a method of distributing them into the five categories shown in Table 3.\* All the 8 linear  $\alpha,\beta$ -unsaturated ketones absorb in the range 1675 to 1685  $\text{cm}^{-1}$  and mass spectrometry can be used to differentiate further within this group. Compounds IV, V, VII and VIII can be identified from their mass spectra as methyl ketones (large  $[M - 15]$  ion) and the *n*-propyl isomer alone of this sub-set affords a McLafferty rearrangement ion at mass 70. The two ethyl ketones X and XI (abundant  $[M - 29]$  ions) and the two vinyl ketones XII and XIII (low intensity but identifiable  $[M - 27]$  ions) can be recognized from their mass spectra, with the latter pair being separable since XIII alone of the two will yield a McLafferty rearrangement ion. A unique identification between the remaining five isomers (Table 3, column 3) can be accomplished by n.m.r. spectroscopy.

The other four major categories in Table 3 are delineated into their respective groups by the position of the carbonyl absorption in the infrared spectrum of each

\* Although we chose infrared absorption spectra to distinguish the five categories shown in Table 3, ultraviolet absorption spectra could have been utilized for an initial distribution of the isomers into sub-classes.

TABLE 3. DIFFERENTIATION BETWEEN ALL KETONE ISOMERS OF  $C_6H_{10}O$  USING I.R., MASS AND N.M.R SPECTROSCOPY


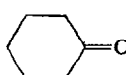
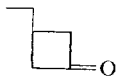


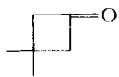
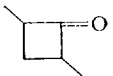
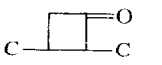
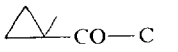
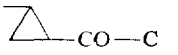
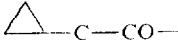
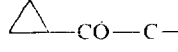
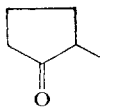
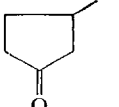
	COMPOUND	MS	NMR
i.r. 1675 to 1685 cm <sup>-1</sup>	$C-CO-C=C-C-C$ (IV)		1 C methyl
	$C-CO-C=C$   C   C (V)	All methyl ketones	2 identical C methyls
	$C-CO-C=C-C$   C (VII)	No McLafferty Rearrangement	2 different C methyls
	$C-CO-C-C-C$    C (VIII)	McLafferty at <i>m/e</i> 70	
	$C-C-CO-C=C-C$ (X)		1 C methyl = doublet 1 C methyl = triplet
	$C-C-CO-C=C$   C (XI)	Ethyl ketones	1 C methyl = singlet 1 C methyl = triplet
	$C-C-CO-C=C$   C (XII)	Vinyl ketone No McLafferty Rearrangement	
	$C=C-CO-C-C-C$ (XIII)	Vinyl ketone McLafferty at <i>m/e</i> 70	
i.r. 1710 to 1720 cm <sup>-1</sup>	$C-CO-C-C-C=C$ (I)		no C methyl
	$C-CO-C-C=C-C$ (II)		1 C methyl = doublet
	$C-CO-C-C=C$   C (VI)	All methyl ketones No McLafferty Rearrangement	no —CH <sub>2</sub> —
	 -CO-C (XXIII)		no vinylic proton
	$C-C-CO-C-C=C$ (IX)	Ethyl ketone	
	$C-CO-C-C=C$   C (III)	Methyl ketone McLafferty at <i>m/e</i> 58	
	 (IV)	Cyclic ketone from m.s.	

TABLE 3 (Contd.)

	Compound	MS	NMR
i.r. 1770 to 1785 cm <sup>-1</sup>	 (XVII)	+D <sub>2</sub> → M + 4	
	 (XVIII)	+D <sub>2</sub> O → M + 3	
	 (XIX)		2 different —CH <sub>2</sub> —
	 (XX)		2 identical —CH <sub>2</sub> —
	 (XXI)		2 identical C-methyl groups
	 (XXII)		2 different C-methyl groups
i.r. 1690 to 1695 cm <sup>-1</sup>	 (XXIV)	All methyl ketones	1 C methyl as a singlet
	 (XXV)		1 C methyl as a doublet
	 (XXVII)		No C methyl signal
	 (XXVI)	Ethyl ketone	
i.r. 1725 to 1740 cm <sup>-1</sup>	 (XV)	+D <sub>2</sub> O → M + 3	
	 (XVI)	+D <sub>2</sub> O → M + 4	

All NMR tests are simple. For instance one C-methyl means a methyl group which is not on a carbonyl function.

isomeric ketone. Further distinction within each of these categories is obtained (Table 3) from either or both the mass and n.m.r. spectral data. In the case of 2- and 3-ethylcyclopentanones (XV and XVI) and again with 3- and 2-ethylcyclobutanone (XVII and XVIII) it would be necessary to rerun the low resolution mass spectrum of each compound following basic exchange in deuterium oxide. Under these conditions the molecular ion of XV and XVIII would increase by 3 mass units and XVI and XVII would increase by 4 mass units.

It is evident from this discussion that the problem of distinguishing between all the isomeric ketones of composition  $C_6H_{10}O$  would be achieved with a computer by programming a specific set of heuristics based on the data summarized in Table 3. This approach is clearly feasible for the 27 ketone isomers of Table 3 where one is concerned with a limited and well defined search space. In those instances (e.g. alcohols (215 cyclic isomers) or ethers (310 cyclic isomers)) where many more compounds of composition  $C_6H_{10}O$  exist, more general heuristics would have to be programmed. This is not as yet feasible since we have very little or no information on the spectroscopy and especially mass spectrometric behavior of the many diversified structures involved. In these instances the problems of first synthesizing many of the yet undescribed compounds, some of which could present a serious synthetic challenge, and the collection of their physical data must be surmounted before extensive programming could be commenced. The magnitude of these problems mitigated against our pursuing this investigation further.

The ability of the computer to generate the individual isomers whose total number are listed in Tables 1 and 2 demonstrates to the organic chemist that many structures of a given composition remain to be synthesized.

Clearly at this stage of development the computer is unsurpassed in the task of generating all, or selected, structures of a given composition. This in itself represents convincing testimony of the power computers will exercise in the future development of organic chemistry.\* One must conclude, however, that at the moment computers are most useful for routine structure determination from physical data in the same specific areas where they already have demonstrated<sup>1</sup> their superiority to the organic chemist.

Before the computer program can successfully solve structure determination problems it is necessary for it to have rules of thumb (heuristics) for differentiating classes of isomers. These rules are currently generated by an organic chemist looking at the mass spectra of representative compounds of the class.† As a matter of interest for this publication all isomers of  $C_6H_{10}O$  were synthesized and their mass spectra studied, although it would seldom be necessary to have the complete list of spectra for the class of compounds under consideration.

#### EXPERIMENTAL

Mass spectra were obtained by Mr R. G. Ross with an MS-9 mass spectrometer operating at 70 eV and an ion source temperature of 180°. Samples were introduced by the heated inlet system.

All compounds were purified by preparative gas chromatography over a 20% carbowax column

\* See for instance E. J. Correy and W. T. Wipke, *Science* **166**, 178 (1969), where a suitably programmed computer has been used to assist in the design of complex organic syntheses.

† The importance of computer learning routines for just this process cannot be underestimated. See for instance P. C. Jurs, B. R. Kowalski, T. L. Isenhour and C. N. Reilley, *Anal. Chem.* **41**, 1949 (1969).

( $\frac{3}{8}$  in.  $\times$  10 ft.). Compounds I, V, VII, XIV, XV, XVI and XXIII were commercially available. Compounds IV, X and XIII were prepared by manganese dioxide oxidation of the corresponding commercially available alcohols. Physical data for compounds XII,<sup>7</sup> XIX,<sup>8</sup> XX,<sup>7</sup> XXI,<sup>7</sup> XXII<sup>7</sup> and XXV<sup>9,10</sup> were acquired from the literature.

4-Hexen-2-one (II)<sup>11</sup> was prepared by acid catalyzed dehydration of 3-hexen-2,5-diol.

Isomesityl oxide (III)<sup>12</sup> was obtained by acid catalyzed enrichment of commercial mesityl oxide.

3-Methyl-4-penten-2-one (VI)<sup>13</sup> and 5-hexen-3-one (IX)<sup>13</sup> were synthesized by reacting dicrotyl and diallyl zinc with ethyl cyanide.

3-Ethyl-3-buten-2-one (VIII)<sup>14</sup> was prepared by the reaction of 2-ethyl-3-ketobutanoic acid with formaldehyde and dimethylamine.

2-Ethylcyclobutanone (XVIII)<sup>15</sup> was obtained by solvolysis of 3-hexyne-1-ol *m*-nitrobenzenesulfonate.

Cyclopropyl ethyl ketone (XXVI)<sup>16</sup> and cyclopropyl acetone (XXVII)<sup>16</sup> were isolated from the reaction of cyclopropyl cyanide and cyclopropyl methyl cyanide with ethyl and methyl magnesium bromides respectively.

4-Methyl-4-penten-3-one (XI). A mixture of ethyl-2-methyl-3-oxovalerate (4.52 g, 28.6 mmole) was stirred at room temperature with 1 N sodium hydroxide (30 ml) for 1 hr. Dimethylamine hydrochloride (2.5 g, 31.0 mmole) and 38% formaldehyde (3 ml) were added, the mixture heated to 85° for 1 hr., acidified with 5 N hydrochloric acid and extracted with ether. The ether extract on evaporation and purification by g.l.c. afforded XI (5% yield). Mass spectrum,  $[M]^+ = m/e$  98; i.r. 1680  $\text{cm}^{-1}$  (carbonyl), n.m.r. in  $\text{CDCl}_3$  ( $\delta$  values) 1.05 (*t*,  $J = 7$  cps), 1.85 (*d*,  $J = 1$  cps, 3H), 2.65 (*q*,  $J = 7$  cps, 2H), 5.68 (*q*,  $J = 1$  cps, 1H), 5.85 (*s*, 1H).

3-Ethylcyclobutyl methyl ketone. Methyl lithium (2.4 g, 0.11 m) in ether was added dropwise to a cooled stirred solution of 3-ethylcyclobutane carboxylic acid<sup>17</sup> (6.4 g, 0.05 m) in ether (10 ml). The mixture was stirred for an additional 30 minutes and decomposed with a saturated ammonium chloride solution. 3-Ethylcyclobutyl methyl ketone (4.5 g, 70% yield) was obtained as a colorless liquid which was homogeneous by g.l.c. Mass spectrum  $[M]^+ = m/e$  126; i.r. 1700  $\text{cm}^{-1}$  (carbonyl); n.m.r. in  $\text{CDCl}_3$  ( $\delta$  values), 0.80 (*t*,  $J = 7.0$  cps, 3H), 1.40 (*c*, 3H), 2.0 (*s*, superimposed on a complex 7H signal), 3.2 (*c*, 1H).

3-Ethylcyclobutyl acetate. This compound was obtained in 81% yield by Baeyer-Villiger oxidation<sup>18</sup> of 3-ethylcyclobutyl methyl ketone. The product was a colorless liquid and the following physical properties were recorded. Mass spectrum,  $[M]^+ = m/e$  142; i.r. 1740  $\text{cm}^{-1}$  (carbonyl); n.m.r. in  $\text{CDCl}_3$  ( $\delta$  values) 0.85 (*t*,  $J = 6.5$  cps, 3H), 1.70 (*c*, 5H), 2.0 (*s*, 3H), 2.58 (*c*, 2H), 4.85 (*c*, 1H).

3-Ethylcyclobutanol. 3-Ethylcyclobutyl acetate (5.0 g, 0.038 m) was hydrolyzed with 10% aqueous potassium hydroxide for 5 hrs. at room temperature. 3-Ethylcyclobutanol was obtained as a colorless liquid (3.35 g, 90% yield) which had the following properties: i.r. 3350  $\text{cm}^{-1}$  (hydroxyl), n.m.r. in  $\text{CDCl}_3$  ( $\delta$  values) 0.85 (*t*,  $J = 6.5$  cps, 3H), 1.40 (*c*, 3H), 2.22 (*c*, 5H).

3-Ethylcyclobutanone (XVII). A solution of sodium dichromate dihydrate (0.50 g, 1.68 mmoles) in sulfuric acid (0.375 ml) and water (2.0 ml) was added dropwise to a vigorously stirred solution of 3-ethylcyclobutanol (0.5 g, 5 mmoles) in ether (8 ml). The mixture was vigorously stirred for 1.5 hrs. After work up the ether layer afforded a pale liquid which contained 55% of 3-ethylcyclobutanone by g.l.c. The following physical properties were recorded for XVII: mass spectrum  $[M]^+ = m/e$  98; i.r. 1780  $\text{cm}^{-1}$  (carbonyl); n.m.r. in  $\text{CDCl}_3$  ( $\delta$  values) 0.92 (*t*,  $J = 6.5$  cps, 3H), 1.60 (*q*,  $J = 6.5$  cps, 3H), 2.50 (*c*, 2H), 3.0 (*c*, 2H).

1-Methylcyclopropyl methyl carbinol. A solution of 3-methyl-3-buten-2-ol (4.3 g, 0.05 m) in anhydrous ether (10 ml) was added dropwise to the organozinc reagent prepared from zinc-copper couple (4.0 g, 0.05 m) and diiodomethane (13.5 g, 0.05 m) in the usual manner. The adduct after standing overnight was decomposed by 1N hydrochloric acid. Distillation of the product isolated from the ether layer after work up and preparative g.l.c. afforded 1-methylcyclopropyl methyl carbinol (0.5, 10% yield) as a colorless liquid. The following physical data were obtained with this product: i.r. 3350  $\text{cm}^{-1}$  (hydroxyl), n.m.r. in  $\text{CDCl}_3$  ( $\delta$  values) 0.40 (*c*, 1.10 (*c*), 1.90 (*c*), 3.20 (*q*,  $J = 6.5$  cps).

1-Methyl-1-acetylcyclopropane (XXIV). 1-Methylcyclopropyl methyl carbinol (240 mg, 2.4 mmoles) was stirred with active manganese dioxide (5 g) in dichloromethane (50 ml) for 5 days. On work up a pale liquid was obtained which on preparative g.l.c. furnished 1-methyl-1-acetyl cyclopropane (XXIV) (70 mg, yield 30%). Mass spectrum  $[M]^+ = m/e$  98; i.r. 1693  $\text{cm}^{-1}$  (carbonyl), n.m.r. in  $\text{CDCl}_3$  ( $\delta$  values) 0.70 (*c*, 2H), 1.20 (*c*, 2H), 1.35 (*s*, 3H), 2.04 (*s*, 3H).



*Acknowledgements*—Financial assistance from the Advanced Research Projects Agency (Contract SD-183), the National Aeronautics and Space Administration (NGR-05-020-004) and the National Institutes of Health (Grants GM 11309 and AM 04257) is gratefully acknowledged.

## REFERENCES

1. Part IV. See A. Buchs, A. M. Duffield, G. Schroll, C. Djerassi, A. B. Delfino, B. G. Buchanan, G. L. Sutherland, E. A. Feigenbaum and J. Lederberg, *J. Am. Chem. Soc.*, **92** (1970).
2. A. M. Duffield, A. V. Robertson, C. Djerassi, B. G. Buchanan, G. L. Sutherland, E. A. Feigenbaum and J. Lederberg, *J. Am. Chem. Soc.* **91**, 2977 (1969).
3. G. Schroll, A. M. Duffield, C. Djerassi, B. G. Buchanan, G. L. Sutherland, E. A. Feigenbaum and J. Lederberg, *J. Am. Chem. Soc.* **91**, 7440 (1969).
4. J. Lederberg, *Proc. Natl. Acad. Sci. U.S.* **53**, 134 (1965); J. Lederberg, 'Topology of Molecules' in *The Mathematical Sciences*, M.I.T. Press, 1969, p. 37.
5. J. Lederberg, *Am. Math. Monthly*, **74**, 522 (1965); J. Lederberg, DENDRAL-64, Part II, 'Topology of Cyclic Graphs', N.A.S.A. CR-68898, STAR No. N66-14074 (1965).
6. J. Lederberg, G. L. Sutherland, B. G. Buchanan, E. A. Feigenbaum, A. V. Robertson, A. M. Duffield and C. Djerassi, *J. Am. Chem. Soc.* **91**, 2973 (1969).
7. J. Kossanyi, *Bull. Soc. Chim. France* 704 (1965).
8. D. S. Weiss, R. B. Gargosian and N. J. Turro, *Org. Mass Spectrom.* **3**, 145 (1970).
9. J. L. Pierre, R. Barlet and P. Arnaud, *Spectrochim. Acta* **23**, 2297 (1967).
10. J. L. Pierre and P. Arnaud, *Bull. Soc. Chim. France* 1690 (1966).
11. H. Morrison, *Tetrahedron Letters* 3653 (1964).
12. R. Mecke and K. Noack, *Chem. Ber.* **93**, 210 (1960).
13. M. Gaudemar, *Bull. Soc. Chim. France* 974 (1962).
14. S. Szántay and J. Rohály, *Chem. Ber.* **96**, 1788 (1963).
15. M. Hanack and I. Herterich, *Tetrahedron Letters* 3847 (1966).
16. M. Hanack and H. M. Ensselin, *Annalen* **697**, 100 (1966).
17. K. C. Brannock, A. Bell, R. D. Burpitt and C. A. Kelly, *J. Org. Chem.* **29**, 801 (1964).
18. W. D. Emmons and G. B. Lucas, *J. Am. Chem. Soc.* **77**, 2287 (1955).